

MENG-ONLINE-KAN ISIS KE WEB DENGAN PEARL DAN HTDIG

Oleh: Hendro Wicaksono
hendrowicaksono@yahoo.com
<http://temubalik.blogspot.com>

Lisensi Dokumen:

Copyright © 2004 Hendro Wicaksono

Seluruh dokumen dapat digunakan, dimodifikasi dan disebarakan secara bebas untuk tujuan pendidikan (nonprofit), dengan syarat tidak menghapus atau merubah atribut penulis dan pernyataan copyright yang disertakan dalam setiap dokumen.

Pendahuluan

CDS/ISIS atau WinISIS (www.unesco.org), selanjutnya disebut ISIS merupakan perangkat lunak berbasis teks yang banyak dipakai untuk pengolahan data bibliografi (pengatalogan) banyak perpustakaan di Indonesia. Seiring dengan makin populernya web sebagai media penyebaran informasi, kebutuhan untuk meng-online-kan ISIS di internet semakin tinggi.

Jika dulu cara meng-online-kan ISIS agak sulit dan biasanya secara teknis agak sulit dipelajari pustakawan. Maka sekarang tersedia banyak cara. Hanya saja meng-online-kan ISIS ke web, merupakan hal yang mudah-mudah susah. Mudah karena begitu banyak solusi alternatif yang bisa digunakan. Susah karena solusi pada satu keadaan belum bisa dipakai pada kondisi yang lain.

Misalnya, meng-online-kan ISIS dengan wwisis atau GreenStone, cocok untuk lingkup intranet (lokal) atau jika anda

punya server yang terkoneksi langsung ke internet. Sehingga anda punya hak untuk meng-install program besar (seperti GreenStone) dan meng-install Operating System kesukaan anda (misalnya Windows untuk wwisis). Masalah akan muncul bila dana dan fasilitas terbatas, yang artinya tidak punya dana sewa hosting dengan space besar atau mempunyai server sendiri dengan *dedicated connection* ke internet.

Akhirnya, solusinya menggunakan software-software yang banyak digunakan web hosting murmer (murah meriah) di internet, yaitu solusi menggunakan PHP (scripting language) dan MySQL (database). Buat sebagian pustakawan, belajar PHP dan MySQL *learning curve*-nya lumayan vertikal. Artinya proses pembelajarannya relatif lebih sulit dan lama.

Sebenarnya ada solusi lain lagi yang bisa digunakan di intranet atau di web hosting murmer dan relatif mudah dipelajari. Salah satunya menggunakan ht://Dig dan Perl. Setiap account web hosting murmer yang

kebanyakan berbasis Linux, biasanya sudah mendapat akses shell dan Perl. Sedangkan ht://Dig harus diinstall dulu, tetapi bisa diinstall tanpa harus punya account root (super user). Keuntungan utama menggunakan ht://Dig adalah hasil penelusuran ditampilkan berdasarkan ranking. Fitur ini tidak terdapat jika meng-online-kan isis menggunakan wwwisis.

Software yang Digunakan

1. Apache Web Server. Silahkan download versi Windows-nya di www.apache.org. Downloadlah format.msi daripada format .exe. Penulis mendownload file `apache_2.0.49-win32-x86-no_ssl.msi` (versi 2.0.49).
2. ht://Dig. Silahkan download versi Windows-nya yang sudah penulis *patch* beberapa file-nya agar *compliant* dengan Apache Web Server di www.geocities.com/hendrowicakso no/htdig-316-w2k.zip.
3. Perl. Silahkan download versi Windowsnya (ActivePerl) di www.activestate.com. Downloadlah format .msi daripada format .exe. Penulis mendownload file `ActivePerl-5.8.3.809-MSWin32-x86.msi`.
4. Bila anda belum mempunyai CDS/ISIS atau WinISIS, silahkan download dari <http://groups.yahoo.com/group/isis-files>.

Untuk Anda ketahui....

Apakah itu .msi? MSI (Microsoft Software Installer) merupakan standar terbaru dari Microsoft yang dibuat untuk memudahkan developer agar

mudah dalam distribusi instalasi software dan menjanjikan integrasi yang lebih baik dengan sistem Windows.

Apa itu ht://Dig? Kalau anda tahu prinsip cara kerja search engine (seperti altavista, Yahoo!, Google), maka anda tidak akan kesulitan dalam memahami cara kerja ht://Dig. ht://Dig adalah suatu sistem pengindeksan dan penelusuran WWW. ht://Dig tidak didesain untuk "bersaing" dengan *powerful search engine* seperti Lycos, Infoseek, Google dan Altavista, tetapi untuk lingkup yang lebih kecil seperti, intranet atau jaringan kampus. Beberapa fitur ht://Dig :

1. Penelusuran Intranet. ht://Dig mempunyai kemampuan untuk mengindeks dan menelusur ke beberapa server sekaligus pada sebuah jaringan dengan "berakting" sebagai browser web.
2. Gratis. ht://Dig dirilis dengan lisensi GNU General Public License.
3. Mendukung Robot Exclusion. Robot Exclusion adalah suatu standar yang digunakan agar bagian-bagian tertentu dari suatu situs web tidak diindeks dan ditelusur oleh Search Engine Robot.
4. Penelusuran Ekspresi Boolean. Mendukung penelusuran kompleks menggunakan logika Boole.
5. Hasil penelusuran yang bisa dikustom. Output penelusuran bisa dibuat sesuai dengan kebutuhan dengan menyediakan template HTML.
6. Fuzzy Searching. Penelusuran bisa dilakukan menggunakan berbagai

- algoritma yang bisa dikustom, seperti: exact, soundex, metaphone, common word endings, sinonim.
7. Menelusur file HTML dan teks. Bisa menelusur dokumen HTML dan file teks.
 8. Penambahan kata kunci (keyword) ke dokumen HTML. Kata kunci bisa ditambahkan ke dokumen HTML dan tidak akan terlihat sewaktu dokumen dibuka.
 9. Peringatan lewat e-mail untuk dokumen yang sudah expired.
 10. Bisa mengindeks server yang diproteksi dengan username dan password.
 11. Dengan menggunakan external parser, ht://Dig juga mampu mengindeks file PDF (Portable Document Format) dan Microsoft Word.

Apa itu Perl? Perl singkatan Practical Extracting and Reporting Language. Sebuah bahasa skripting buatan Larry Wall, seorang ahli computer linguistics. Perl terkenal karena kemampuannya yang luar biasa untuk pemrosesan teks. Tidak heran Perl menjadi *tool* yang populer untuk system administrator dan bahasa yang banyak digunakan untuk pemrograman CGI (Common Gateway Interface). Dalam artikel ini diuraikan secara ringkas mengenai teknik, Perl yang digunakan untuk mengkonversi data ISIS dari satu file untuk seluruh record menjadi satu file per record.

Dalam artikel ini, tidak diberikan contoh instalasi ht://Dig pada platform Unix/Linux, anda bisa melihat tutorial yang penulis buat di <http://www.pandu.org/Solusi/Htdig/> untuk instalasi ht://Dig di Unix FreeBSD.

OS yang digunakan disini adalah Windows 2000 Advanced Server, dengan pertimbangan Windows paling banyak digunakan oleh pustakawan di Indonesia. Apabila anda menerapkan pada web hosting berbasis Unix/Linux, perbedaannya Cuma pada cara instalasi ht://Dig.

Proses meng-online-kan database ISIS dengan ht://Dig adalah:

1. Data ISIS diprint ke file dengan format HTML. Semua record diletakkan dalam satu file.
2. Dengan Perl, dari satu file untuk semua record, dibuat menjadi satu file per record.
3. File-file tersebut dicopy ke web server.
4. File-file tersebut diindeks oleh ht://Dig.
5. Kemudian katalog bisa di search via ht://Dig.

Temukan petunjuk instalasi Dig pada tutorial yang sudah dibuat dan diletakkan di web seperti ditunjukkan pada alamat web diatas..

Konversi Data ISIS

Penulis menggunakan CDS/ISIS versi 3.07. Menurut penulis, CDS/ISIS versi DOS relatif lebih stabil dan tidak banyak masalah dalam instalasi dan penggunaannya diberbagai platform windows terbaru. Sedangkan database yang diujicoba dalam artikel ini adalah **CDS** yang biasanya terdapat dalam setiap instalasi CDS/ISIS.

Sebelum mulai melakukan konversi data, perlu dirancang dahulu, template HTML

yang akan digunakan untuk menampilkan data ISIS.

Setelah menentukan template HTML, berikutnya adalah membuat Print Format (PFT file) pada CDS/ISIS.

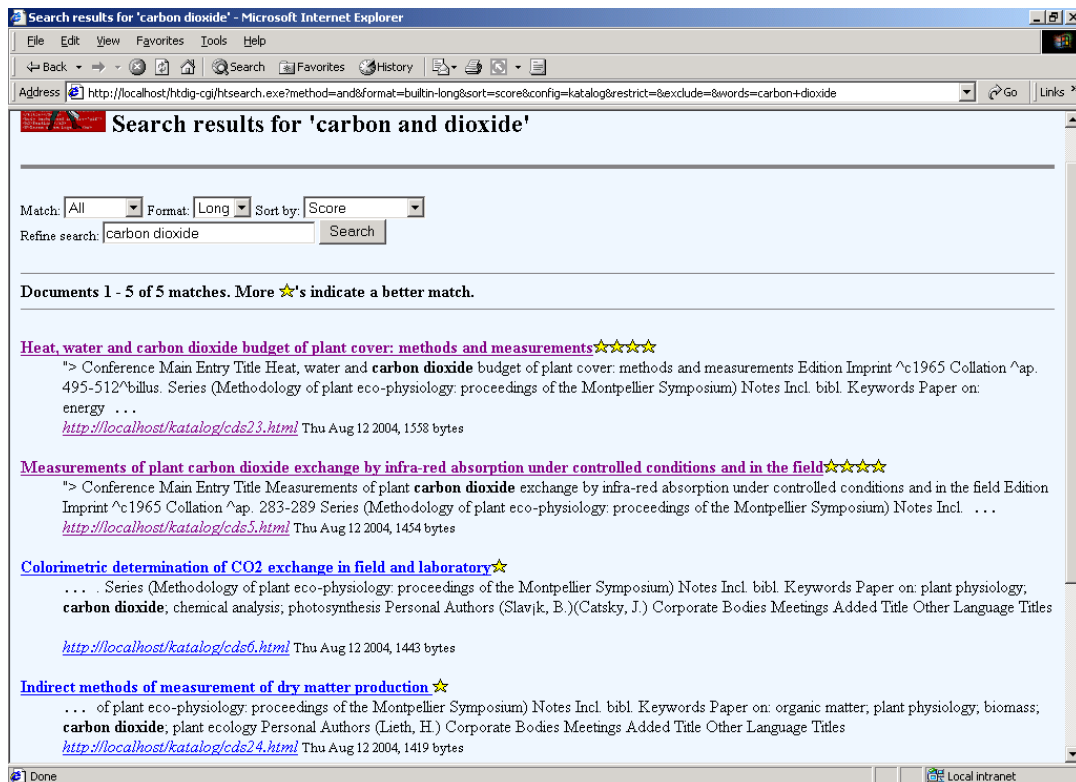
Anda ketahui.....

*Kalau anda tidak senang menggunakan interface CDS/ISIS untuk meng-edit print format (karena tampilannya menjadi kriptik), anda bisa menggunakan Notepad. Caranya, edit saja file **CDS.PFT** yang terdapat pada direktori c:\isis\data. Dengan begini, tampilannya jadi lebih enak dilihat dan bug tracking nya jadi lebih mudah.*

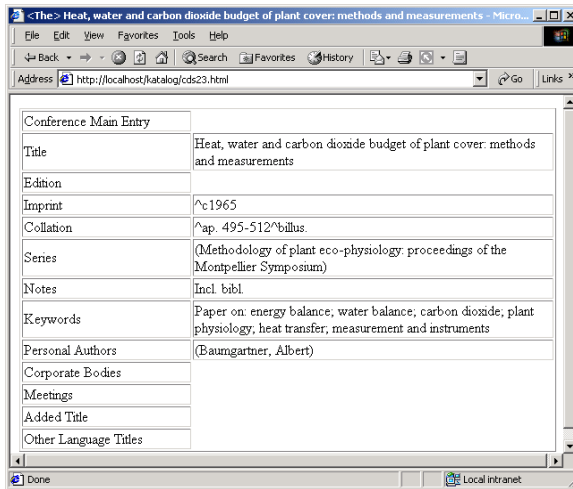
Berikut ini cara melakukan print ke file di CDS/ISIS.

Kemudian akan ada file bernama **TOHTML** yang terletak pada direktori **c:\isis**. Tapi letak ini tidak selalu sama, tergantung dengan setting **syspar.par** instalasi CDS/ISIS anda. File **TOHTML** berisi semua record dari database **CDS**. Berikutnya adalah membuat file-file baru, dimana tiap record terletak dalam satu file. Untuk ini kita akan membuat sebuah skrip Perl sederhana yang fungsinya memindahkan tiap record menjadi file tersendiri.

Dengan browser, akses <http://localhost/htdig/search.html>. Penulis mencoba melakukan penelusuran dengan Keyword "**carbon dioxide**".

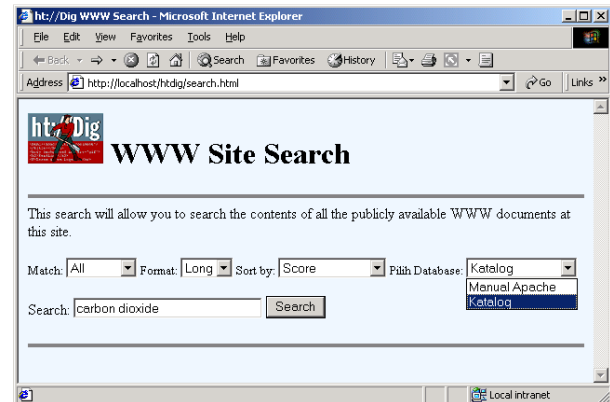


Hasil penelusuran ditampilkan berdasarkan ranking dengan lambang bintang sebagai indikator. Semakin banyak bintangnya berarti rankingnya semakin tinggi atau semakin dianggap tepat.



Tampilan detail katalog. Untuk memperbaiki tampilan katalog, misalnya untuk menghilangkan karakter “^” atau menghilangkan field yang tidak ada isinya, anda bisa meng-edit lebih lanjut file CDS.PFT. Silahkan merujuk ke banyak referensi yang membahas print format di ISIS lebih mendalam.

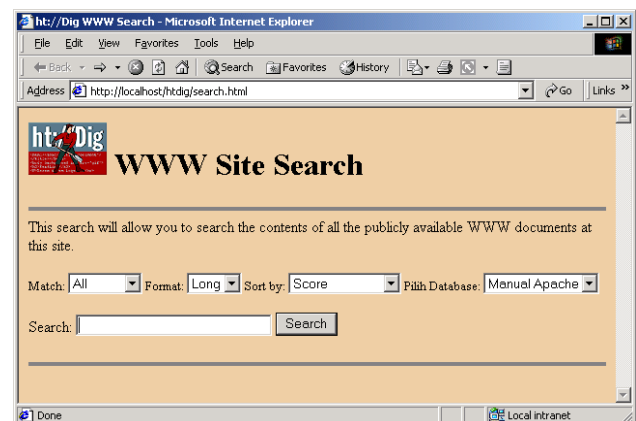
Bagaimana Caranya Memberikan pilihan Database sewaktu memulai penelusuran tanpa harus bolak-balik meng-edit file c:\htdig\html\search.html?



Tampilan penelusuran akan menjadi seperti ini.

Bagaimana Caranya Mengubah Template Untuk diCustomise?

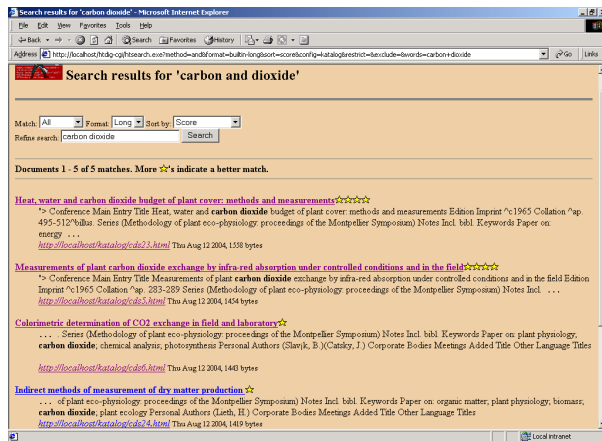
Untuk Tampilan awal penelusuran, file yang harus diedit adalah file c:\htdig\html\search.html. Untuk contoh sederhana, penulis akan mencoba merubah background warna dari eef7ff menjadi efcca3.



Tampilannya akan menjadi seperti ini.

Berikutnya adalah tampilan pada saat ht://Dig menampilkan hasil penelusuran. File-file yang terkait dengan template hasil penelusuran terdapat di direktori:

c:\htdig\common. Untuk mengubah warna background template hasil penelusuran file yang harus diedit adalah **header.html**. Pada file tersebut ubah nilai: **<body bgcolor="#eef7ff">** menjadi **<body bgcolor="#efcca3">**.



Warna background akan berubah menjadi seperti gambar.

Untuk customise lebih lanjut silahkan ujicoba sendiri file-file yang terdapat pada direktori c:\htdig\html dan c:\htdig\common.

Penutup

Ht://Dig sebenarnya punya banyak kelebihan lain, diantaranya juga bisa digunakan untuk membangun digital library skala menengah, karena kemampuannya meng-indeks file-file dengan format populer seperti MS Word dan PDF. Selain itu, ht://Dig juga bisa digunakan untuk meng-indeks data dalam database, sehingga pencarian bisa lebih baik dan berbasis ranking, tidak hanya kronologis.

Referensi

<http://www.htdig.org>

Permadi, Agus; Priyocahyono, Eko S.
*CDS/ISIS Untuk Pustakawan,
Dokumentalis dan Arsiparis*. Jakarta:
Yasmin, 1996

Wicaksono, Hendro. *Membangun Search
Engine Dengan ht://Dig*.
<http://www.pandu.org/Solusi/htdig>